Subterranean Biology 10: 25-35, 2012 (2013)
doi: 10.3897/subtbiol.10.4154

Molecular divergence and evolutionary relationships among Aemodogryllinae from Southern China, Laos and Thailand (Orthoptera, Rhaphidophoridae)

Valerio KETMAIER [1,2,*], Claudio DI RUSSO [1], Mauro RAMPINI [1],
Nadine BERNHARDT [2], Ralph TIEDEMANN [2], Marina COBOLLI [1]

[1] Dipartimento di Biologia e Biotecnologie "Charles Darwin", Università di Roma "La Sapienza", V.le dell'Università 32, I-00185 Roma, Italy
[2] Unit of Evolutionary Biology/Systematic Zoology, Institute of Biochemistry and Biology, University of Potsdam, Karl-Liebknecht-Strasse 24-25, Haus 25, D-14476, Potsdam, Germany; email: ketmaier@uni-potsdam.de
*Corresponding author: ketmaier@uni-potsdam.de

ABSTRACT
In this study we screened for sequence polymorphisms at one mitochondrial (Cytochrome Oxidase subunit I) and one nuclear (Internal Transcribed Spacer 1) gene 33 populations of the cave cricket genera *Diestrammena, Paradiestrammena*, *Eutachycines* and *Paratachycines* from Southern China (three Provinces: Jiangxi, Guangdong and Guizhou), Laos and Thailand. Twenty-five of these populations were assigned to the genus *Diestrammena*, subgenus *Gymnaeta*, while the remaining eight belonged to the genera *Paradiestrammena* (3), *Eutachycines* (3) and *Paratachycines* (2). The degree of troglomorphosis varies among them; some populations are blind and depigmented, some have fully developed eyes, while some others show intermediate characteristics. Phylogenetic searches carried out on the two gene partitions separately revealed multiple cases of incongruence but only three of them were statistically significant and were hence removed from the subsequent analyses based on the combined data set. Our data do not support *Diestrammena* as monophyletic while representatives of *Paradiestrammena*, *Eutachycines* and *Paratachycines* were clustered together; the validity of some nominal species was confirmed molecularly but we also revealed a large number of deeply divergent lineages. Populations with the same degree of troglomorphosis do not cluster together. We identified five major clades; divergence among them (and in a few circumstances also within them) is always higher than the DNA barcode threshold for intraspecific comparisons in insects. In two circumstances, the same clades (III and V) are co-distributed in geographically distinct areas (Provinces). This geographical distribution might be explained by envisioning an evolutionary scenario based on zones of secondary admixture following epigean dispersal among lineages that diverged in allopatry.

Key words: Aemodogryllinae, cave cricket, mitochondrial DNA, nuclear DNA, secondary contact.

INTRODUCTION

Southern China, with about 2.6 million km^2, hosts one of the largest karstic areas of the world and a vast – yet largely unexplored – biodiversity. Starting from the early '90s, a number of biospeleological expeditions have been organized in the area (Latella and Zorzin 2008). These were joint efforts carried out by Chinese, French, Italian and Slovenian teams. Even though these expeditions could visit only a small fraction of the caves in the area, the amount of speleological and biological information they gathered is massive and still in the course of being processed. Among others, cave crickets belonging to the subfamily Aemodogryllinae (Rhaphidophoridae) were retrieved in quite a large number (Rampini et al 2008).

The family is distributed worldwide and includes epigean representatives (mostly living in the litter of tropical forests) and many subterranean forms confined mostly to temperate areas where conditions for epigean life are favorable only seasonally. In the Mediterranean area two genera are present (*Dolichopoda* and *Troglophilus*); both have been object of extensive evolutionary research (Allegrucci et al 2011; Ketmaier et al 2010 and references therein). Allegrucci et al (2010) produced a biogeographic reconstruction of the evolutionary relationships among Rhaphidophorids from the Southern End of the World.

Here, we took advantage of the cave crickets collected in Southern China in three Provinces (Jiangxi, Guangdong and Guizhou) to assess their degree of genetic divergence and to produce a preliminary phylogeographic hypothesis on the diversification of the group in the area. We additionally included in our study cave crickets collected in Laos and Thailand. A detailed morphological assessment of these organisms (Rampini et al 2008) revealed that most of the samples at our disposal belong to the genus *Diestrammena*. In particular, Southern Chinese populations could be readily identified as belonging to the subgenus *Gymnaeta* because of the genital morphology and because they lack spines on the hind femurs; samples from Laos and Thailand were attributed to the genera *Eutachycines, Paradiestrammena* and *Paratachycines* (Storozhenko 1991; Gorochov 1994; Gorochov 2002). Some populations could be identified at the species level (see Table 1); other populations were taxa new to science while

Received 22 October 2012; Accepted 15 December 2012

some others are still awaiting a formal specific attribution (Rampini et al 2008). All these taxa show a varying degree of adaptation to life in caves (troglomorphosis); some are blind and with reduced body pigments, some have reduced eyes while others have fully developed eyes and a pigmented body. We analyzed sequence polymorphisms at partial sequences of two genes, the mitochondrial (mtDNA) Cytochrome Oxidase subunit I (COI) gene and the nuclear (nDNA) Internal Transcribed Spacer 1 (ITS1). These markers have already proved useful in unveiling evolutionary relationships at the taxonomic level tackled in the present study in the cave cricket genus *Troglophilus* (Ketmaier et al 2010). We specifically ask whether the different genera could be validated molecularly, whether populations with the same degree of troglomorphosis cluster together and whether populations from the same geographic area are also genetically close.

MATERIALS AND METHODS

Sampling

Study samples were collected from twenty-nine caves from three South Chinese Provinces (Guizhou, Guangdong and Jiangxi) and from Thailand and Laos. When feasible, sampled individuals were assigned to nominal species following Gorochov et al (2006) and Rampini et al (2008). In two cases we found the occurrence of different species or genera in the same cave (i. e. *Diestrammena chenui* and *Diestrammena ferecaeca* in An Ja Da Dong from Guizhou, and *Eutachycines cassani* and *Paradiestrammena vernalis* in the Maria Cassan cave from Laos). Furthermore, in two caves from Guangdong (Pi Pa Dong and Xian Yen Dong) we found individuals with different degrees of troglomorphosis; we kept these individuals distinct (i.e. not lumped in a single population) in the statistical treatment of data. At total of 93 individuals were collected and analyzed genetically. Details on the sampled localities and populations can be found in Table 1 and Fig. 1. Samples were collected by and preserved in 70% ethanol; part of the caudal femur muscle was used for the genetic work. Analyzed individuals are deposited in the M. Rampini collection (Department of Biology and Biotechnology "C. Darwin"; University of Rome "Sapienza", Rome, Italy).

Genetics

Total genomic DNA was extracted following Ketmaier et al (2008). We amplified by Polymerase Chain

Table 1 - Populations and species (whenever a specific attribution was feasible) of Aemodogryllinae analyzed molecularly. For each taxon we indicate the Southern Chinese province and cave of origin and the population code used in the study. The last column shows the sample size (N). Underlined caves are those where two genetically divergent lineages co-exist.

Taxon	Province*	Cave	Code**	N
D. (Gymnaeta) sp.	Jiangxi	Hengli Dong	C23c	2
D. (Gymnaeta) sp.	Guangdong	Tan Shui Long Yan Dong	C10b	6
D. (Gymnaeta) sp.	Guangdong	Wei Yan Dong	C12b	3
D. (Gymnaeta) sp.	Guangdong	Jing Long Dong	C21b	3
D. (Gymnaeta) sp.	Guangdong	Shi Shi Qiao Dong	C16b	15
D. (Gymnaeta) sp.	Guangdong	Pi Pa Dong	C17b	5
D. (Gymnaeta) sp.	Guangdong	Ting Xia Wo Dong	C14b	9
D. (Gymnaeta) sp.	Guangdong	Xian Yen Dong	C19b	4
D. (Gymnaeta) sp.	Guangdong	Chao Yan Dong	C9b	1
D. (Gymnaeta) sp.	Guangdong	Da Qiao Dong	C22a	3
D. (Gymnaeta) sp.	Guangdong	Pi Pa Dong	C18a	2
D. (Gymnaeta) sp.	Guangdong	Xian Yen Dong	C20a	3
D. (Gymnaeta) sp.	Guizhou	Yun Dong	C1b	1
D. (Gymnaeta) aspes	Guizhou	Di E Dong	B1.0b	2
D. (Gymnaeta) aspes	Guizhou	Wangma Dong	B1.1b	3
D. (Gymnaeta) chenhui	Guizhou	An Ja Da Dong	B6.1b	1
D. (Gymnaeta) latellai	Guizhou	Shachang Dong	C8b	1
D. (Gymnaeta) ferecaeca	Guizhou	Xiao Dong	B4.8.2a	1
D. (Gymnaeta) ferecaeca	Guizhou	Tiao Shuz Dong	B4.17a	9
D. (Gymnaeta) sp.	Guizhou	Tu Shi Dong	C4a	1
D. (Gymnaeta) sp.	Guizhou	Xiao Dong	C11a	1
D. (Gymnaeta) sp.	Guizhou	Yin Jia Dong	C5a	1
D. (Gymnaeta) ferecaeca	Guizhou	Tang Bian Da Dong	B1.2a	2
D. (Gymnaeta) ferecaeca	Guizhou	An Ja Da Dong	B5.0.1c	5
D. (Gymnaeta) sp.	Laos	Tham Kang	S4a	1
Eutachycines sp.	Laos	Mahaxay Cave	S12a	1
Eutachycines cassani	Laos	Maria Cassan Cave	S15c	1
Eutachycines cassani	Laos	Tham Nja Kay Khia	S6c	1
Paradiestrammena vernalis	Laos	Ban Tham	S9b	1

Taxon	Province*	Cave	Code**	N
Paradiestrammena vernalis	Laos	Marian Cassan Cave	S11b	1
Paradiestrammena vernalis	Laos	Tham Deua	S13b	1
Paratachycines thailandensis	Thailand	Tam Lod	S1b	1
Paratachycines thailandensis	Thailand	Christmas Cave	S2a	1

* Southern Chinese provinces only; it does not apply to populations from Laos and Thailand.

** a = small eyes; b = normal eyes; c = eyes absent.

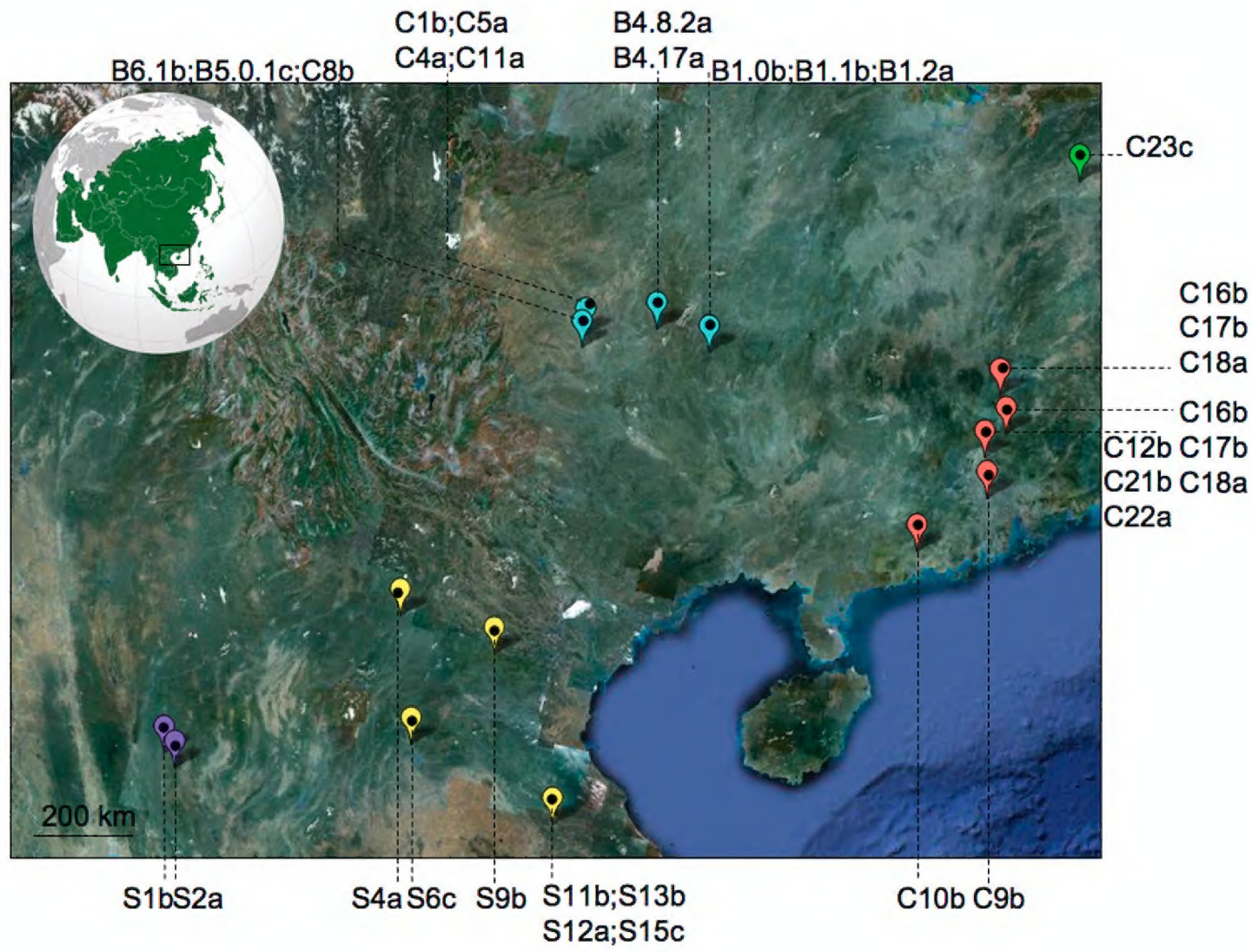

Fig. 1 - Geographic distribution of the localities sampled for the study. Population codes are as in Table 1. Green is Jiangxi, red is Guangdong, light blue is Guizhou (Southern Chinese Provinces) while Laos and Thailand are indicated in yellow and violet, respectively.

Reaction (PCR) a fragment of the mitochondrial DNA (mtDNA) region encoding for the Cytochrome Oxidase subunit I gene (COI) and a fragment of the nuclear DNA (nDNA) Internal Transcribed Spacer 1 (ITS1). For COI we used the same set of primers as in Ketmaier et al (2008) while for the ITS1 gene we adopted the primer pair Cas5p8sB1d/ Cas18sF1 developed by Ji et al (2003). Double-stranded PCR amplifications were performed in a 50 μL reaction volume containing 10mM Tris-HCl (pH 8.8), 50 mM KCl, 1.5 mM MgCl2, each dNTP at 2.5 mM, each primer at 1 mM, genomic DNA (10-100 ng) and 5 units of Amplitaq (Applied Biosystems) following the conditions detailed in Ketmaier et al (2010). For some samples with a low quality of the total DNA, the COI and ITS1 fragments were obtained using a nested PCR approach. For the COI gene we used the primers UEA1 and UEA10 (Lunt et al 1996) for the first PCR round and the same primer pairs as above as internal primers for the nested PCR. The primer pair Cas5p8sB1d/ Cas18sF1 were used for the initial PCR for the ITS1 gene while we designed ex-novo an internal primer pair (ITS-f: 5'- GCGGTCGCTGA TGTCA-3'; ITS-r: 5'- CT-GCAGTTCACATGTCGA-3') for the nested PCR. In either case, PCR reactions were prepared as detailed above but with a different thermal profile (initial denaturation at 98° for 30 sec followed by 30 cycles each consisting of a denaturation step at 98° for 10 sec, annealing at 50° for 20 sec, extension at 72° for 30 sec followed by a final elongation at 72° for 5 min). PCR products were purified using the NucleoSpin® Extract II (Macherey-Nagel).

Purified PCR products were directly sequenced on an automated ABI 3100 sequencer (Applied Biosystems) following the manufactures' protocols. To promote accuracy, each PCR product was sequenced in both directions.

Sequences were edited with Sequencher 4.6 (Gene Code Corporation, Ann Arbor, MI). COI sequences were aligned by eye following the guide provided by the reading frame; for the ITS1 gene we used ClustalX (Thompson et al 1997) with default parameters. Minor adjustments to the alignment were done by eye.

Haplotypes (both genes combined) were identified using the program TCS (Clement et al 2000). We used PAUP* 4.0b10 to calculate base frequencies and to test for base frequency homogeneity across taxa (χ^2 test); PAUP was also used to calculate COI-based Kimura 2-parameter (K2P) distance values among and within the main clades identified phylogenetically. The occurrence of saturation of sequences in the data set was visually evaluated by plotting the absolute number of transitions (Ti) and transversions (Tv) against the uncorrected-*p* distance values. This analysis was performed on four different data partitions: the two genes separately, the two genes combined and, for the COI gene, on 3rd codon positions only. The two data sets were analyzed phylogenetically with the Southern Italian Rhaphidophorid species *Troglophilus andreinii* as outgroup (Ketmaier et al 2002) by maximum parsimony (MP; heuristic searches, ACCTRAN character-state optimization, 100 random step-wise additions, TBR branch-swapping algorithm) (Farris 1970), maximum likelihood (ML; heuristic searches, 100 random stepwise additions, TBR branch swapping algorithm) (Felsenstein 1981), Neighbor-Joining (NJ) (Saitou and Nei 1987) and Bayesian methods (Rannala and Yang 1996; Mau and Newton 1997; Larget and Simon 1999; Mau et al 1999; Huelsenbeck 2000). MP, ML and NJ analyses were performed using PAUP* 4.0b10 (Swofford 2002); Bayesian analysis was carried out using MRBAYES 3.1 (Ronquist and Huelsenbeck 2003). MP searches were run giving equal weight to all substitutions. We ran the ML analyses on PAUP* 4.0B10 after having determined the best model of DNA substitutions that fit our data using MODELTEST (Posada and Crandall 1998). MODELTEST selected the GTR + I + G model as the best fitting for both the COI and ITS1 data partitions (shape parameters $\alpha = 0.454$ and 1.993 for COI and ITS1, respectively). We used this model with the partition-specific settings to run all our ML analyses. NJ analyses were carried out on ML distances calculated with the same settings used for the ML analyses. For the Bayesian approach, we employed the same model of sequence evolution as in the ML searches; for COI we allowed site-specific rate variation partitioned by codon positions. MRBAYES was run for 2 million generations with a sampling frequency of 100 generations. We ran one cold and three heated Markov chains. From the 20000 trees found, we discarded the first 10% ("burn- in") in order to include only trees for which convergence of the Markov chain had been reached.

We followed the approach detailed in van der Niet et al (2008) to identify incongruence between the two data sets. We initially produced strict consensus topologies from separate parsimony analyses based on either data set and inspected them visually. A haplotype (or clade) was identified as a source of incongruence if its removal solved the incongruence. At each step, we removed from the alignment a single identified source of incongruence and we produced a new, pruned, alignment. This procedure was reiterated until no incongruence was left. We then started adding these incongruent haplotypes or clades individually to the pruned combined alignment. Congruence between COI and ITS1 partitions in the alignment was tested with the Length Difference (ILD) test (Farris et al. 1994) in PAUP* 4.0B10. We ran the ILD test after having removed the outgroups and non-informative characters, with 1000 randomizations and 50 stepwise random addition sequence replicates (RASR). If the ILD test showed that adding a given haplotype (or clade) was not the source of a significant incongruence ($P_{ILD} < 0.05$) we kept it in the alignment, otherwise we removed it again. We then assessed whether conflicting positions of incongruent sequences were robust in both data sets separately by running Bayesian searches with the parameters and settings described above. The placements of incongruence were assumed to be robust if supported by posterior probability (PP) values of 75% or greater in both Bayesian analyses. We then plotted the PP values supporting conflicting placements (both genes separately) against the *P*-value of the ILD test. In doing so we were hence able to identify, for each gene, the critical bootstrap value at which the ILD test returns a significant result and, ultimately, to distinguish non-complex from complex cases of incongruence. The complex cases of incongruence were removed from the final alignment and phylogenetic searches run with the same methods as detailed for the two genes separately. The ML and Bayesian searches were based on the re-estimated GTR + I + G model of sequence (in MODELTEST) with a shape parameter of $\alpha = 1.545$. The competing placements of complex cases of incongruence were added to the obtained phylogentic hypothesis by hand.

RESULTS

Sequence variation

For each of the 93 individuals included in the study we sequenced 1246 base pair (bp) of DNA; of these 543 bp belonged to the mtDNA COI gene and 703 bp to the nDNA ITS1 gene. Electropherograms resulting from direct sequencing of the ITS1 gene showed no evidence of

ambiguous base callings rendering cloning unnecessary. Table 2 shows the base frequencies of the different data partitions and the result of χ^2 tests for base frequency homogeneity across taxa. The four bases have very similar frequencies for the ITS1 gene, while COI sequences are anti-G biased and A+T rich especially in third codon positions, as expected for a protein coding mitochondrial gene. No data partition violated the homogeneity of base frequencies across taxa ($0.512 < P\chi 2 < 1.000$).

Fig. 2 shows the saturation plots for the two genes separately and combined and for COI 3rd codon positions. Transversions tend to increase linearly with the increase of genetic divergence whereas a slight saturation of transitions is evident at the higher levels of genetic divergence especially for COI 3rd codon positions.

Sequences of the two genes combined defined a total of 37 unique haplotypes for the 93 individuals included in the study; the absolute frequency of these haplotypes in the different populations is shown in Table 3. Very few haplotypes are shared among different populations while six populations present more than one haplotype. Haplotype 5 is noteworthy in that three individuals from two provinces (B4.8.2a from Guizhou; C12b and C21B from Guangdong) with different degrees of troglomorphosis (B4.8.2a has small eyes while C12b and C21B have both fully developed eyes) share it.

Table 2 - Base frequencies and results of the χ^2 tests for base frequency homogeneity across taxa in the two genes analyzed in the study. Values are shown for COI and ITS1 separately and combined; for COI analyses were conducted on each codon position (1st, 2nd and 3rd) and on all codon positions (all).

Partition	Base frequencies (%)				χ^2 values	$P\chi^2$
	A	C	G	T		
ITS1	25.1	24.1	25.5	25.3	185.8	0.999
COI- 1st	28.7	17.2	26.4	27.7	50.1	1.000
COI-2nd	20.3	22.9	17.0	39.9	19.3	1.000
COI- 3rd	46.9	12.4	5.4	35.4	283.6	0.512
COI- all	32.0	17.5	16.2	34.3	144.4	1.000
ITS1+COI	28.6	20.8	20.8	29.8	203.3	0.999

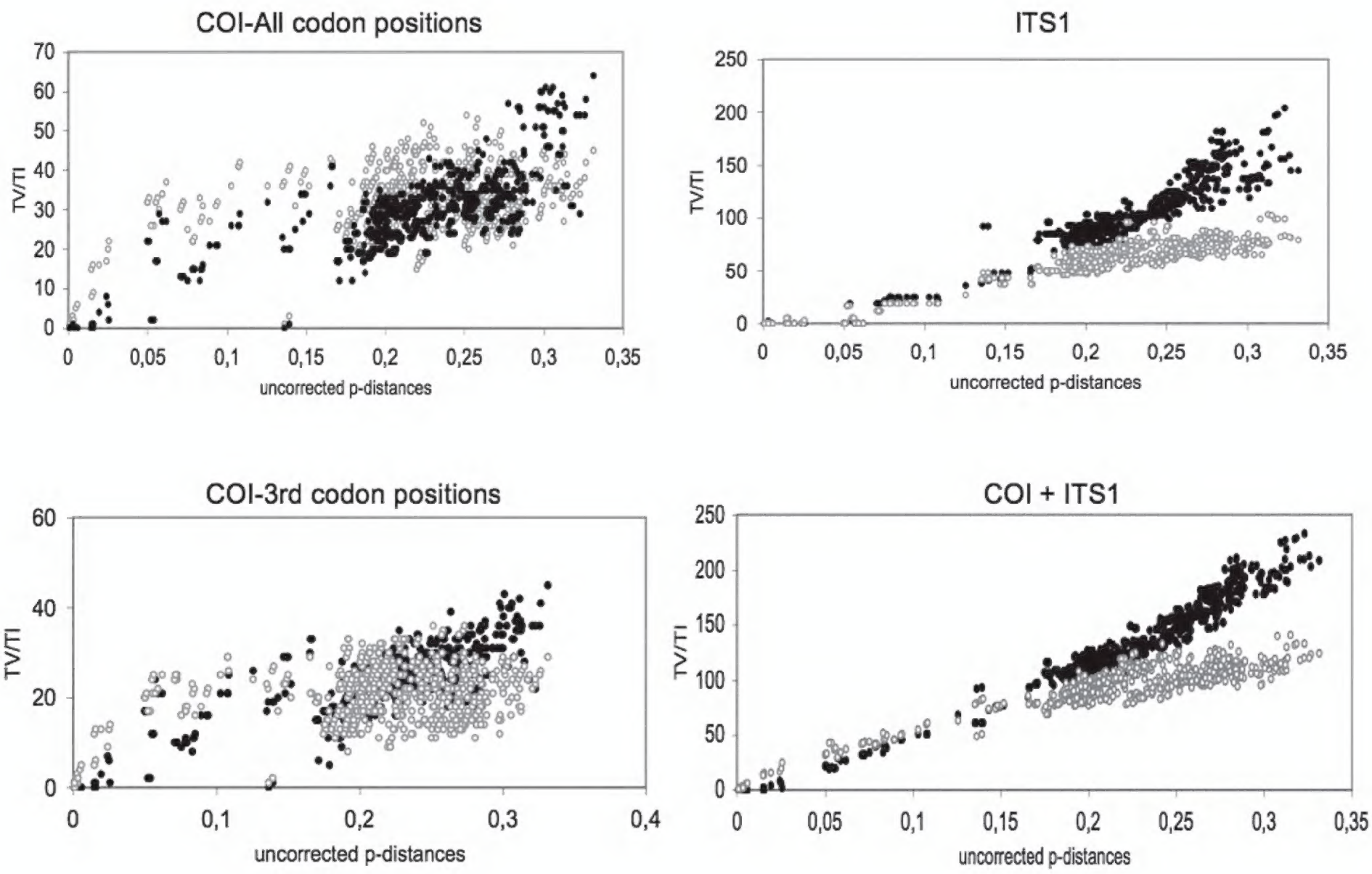

Fig. 2 - Plots of the uncorrected p values against the observed number of transitions (white dots) and transversions (black dots) on all pairs of sequences on each gene separately and both genes combined. For the COI gene we analyzed all codon positions together and 3rd codon positions only.

Table 3 - Absolute haplotype frequencies in the populations and species included in the study. Population codes are as in Table 1.

	Haplotypes																																		
Pop.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37
C1b																			1																		
B1.0b	2																																				
B1.1b		2	1																																		
B6.1b								1																													
C8b																											1										
C10b									4	2																											
C12b					3																																
C21b					1																	2															
C16b															10	3	2																				
C17b															1	3	1																				
C14b												5	3	1																							
C19b												3	1																								
S1b																													1								
S9b																																					1
S11b																														1							
S13b																																1					
C9b																												1								
B4.8.2a					1																																
B17a							9																														
C4a																									1												
C11a											1																										
C5a																										1											
B1.2a				2																																	
C22a																							3														
C18a																		2																			
C20a																				2	1																
S2a																																		1			
S4a																																			1		
S12a																															1						
S15a																																	1				
S6c																																			1		
B5.0.1c						5																															
C23c																								2													

Sources of incongruence and phylogenetic relationships

Plotting of the *P*-value returned by the ILD test against the critical PP value (75%) revealed three complex sources of incongruence (Fig. 3), which corresponded to the populations B1.2a, B4.17a and C1b. All the other non-congruent placements ($P_{ILD} < 0.05$) were not supported by the corresponding PP value (PP < 75%; i.e. non-complex incongruent taxa). It should be noted here that since a PP value of 75% is considered a weak support to relationships in a Bayesian framework (Alfaro et al. 2003), our approach is quite conservative in identifying the main sources of incongruence between the two markers.

Fig. 4 shows the Bayesian tree based on the combined data set pruned of the three complex incongruent populations and summarizes the result of the MP and NJ searches. The Bayesian, MP and NJ topologies were largely similar to one another; incongruent taxa were placed in the position that they held in the topologies resulting from separate analyses of COI and ITS dataset (not shown). Representatives of the genus *Diestrammena* (subgenus *Gymnaeta*) do not cluster in a monophyletic clade since *Eutachycines*, *Paratachycines* and *Paradiestrammena* are deeply embedded within it. Similarly, haplotypes are not grouped in the phylogenetic tree according to the degree of eye development. Two (B5.0.1.c and C23c) of four blind taxa are rather basal in the tree while S6c and S15c are embedded in clade clade IV (see below); many supported clades repeatedly cluster together populations with fully developed and reduced eyes. Five major clades (labeled I-V in Fig. 4) could be identified phylogenetically. Clades I and II are restricted to the Guangdong Province only. Conversely, clades III and V include the Guangdong and Guizhou Provinces, clade IV is spread across locations from Thailand and Laos and includes *Eutachycines, Paratachycines* and *Paradiestrammena*. Finally, clade V covers two provinces but at the within-clade level populations do not necessarily group according to their geographic proximity. Average COI-based K2P divergence among clades ranges between 7.9% (SD = 2%; clade I vs. clade II) and 13.7% (SD = 1.9%; clade IV vs. clade V); at the within-clade level K2P is the lowest within clade II (3%; SD = 1%) and the highest within clade IV (12.6%; SD = 4.3%).

DISCUSSION

We detected three complex sources of incongruence between the COI and the ITS data sets; while this figure is relatively little yet discrepancies between different gene trees should be evaluated carefully when attempting to produce a species tree by combing different genes. Gene and species trees tend to be in good agreement with each other when population sizes are relatively small relative to the branch lengths of a given phylogeny. Our tree of Fig. 4 contains many clades with relatively long branches; no information exists on the demography of these cave crickets but populations of subterranean organisms are seldom exceedingly large and this could explain why only three lineages show irreconcilable placements in the COI and ITS trees. These could be due to the fact that the common ancestry of gene copies extends deeper in time that the splitting event(s). Another explanation could invoke saturation of substitutions especially in the COI partition

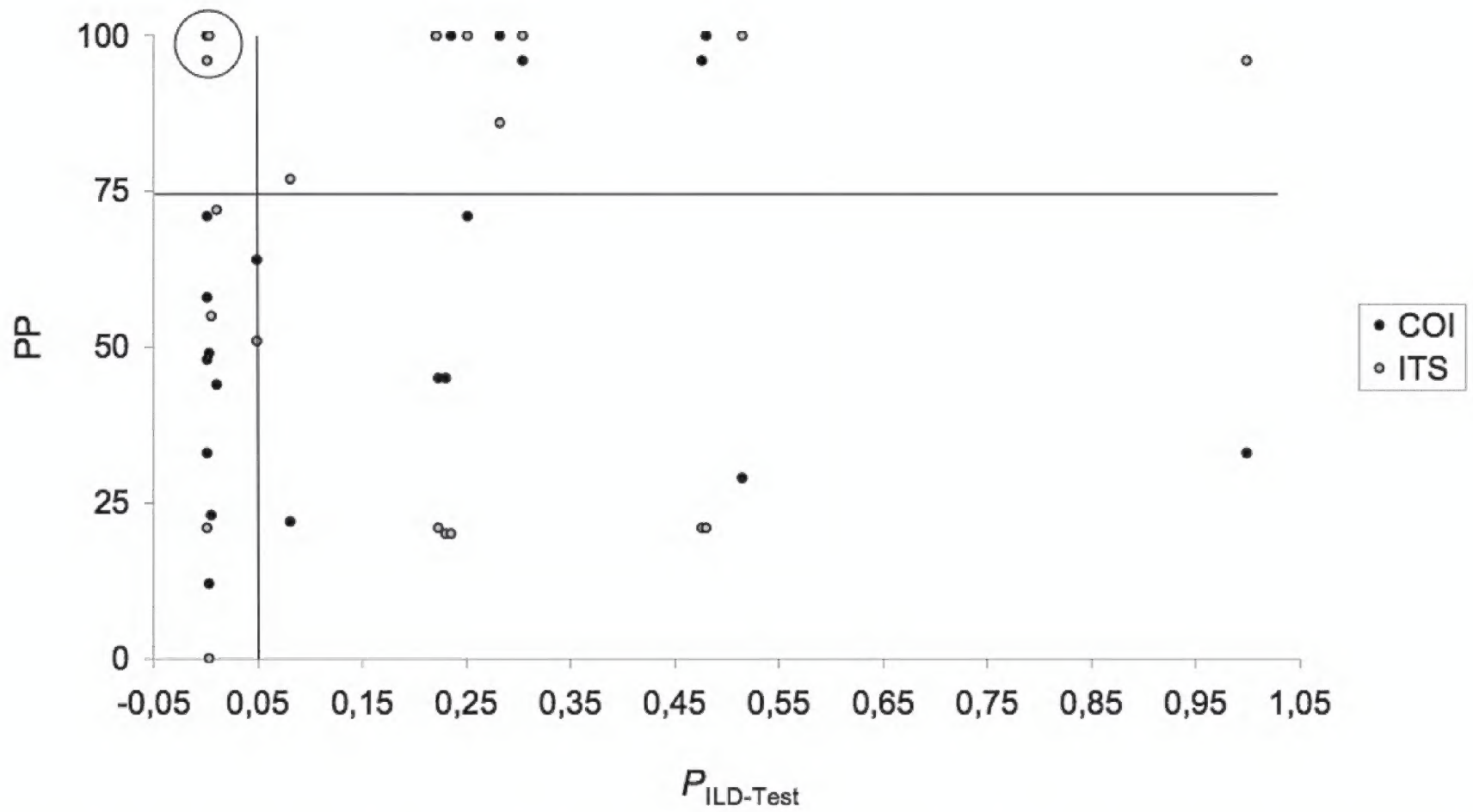

Fig. 3 - Relationship between the critical Posterior Probability (PP) value (75%) and the *P*-values returned from the ILD-test in the two gene partitions. Complex cases of incongruence are those above the PP threshold and with a $P_{ILD_test} < 0.05$ (circled dots). The horizontal line is at 75% PP, the vertical line at $P_{ILD_test} = 0.05$.

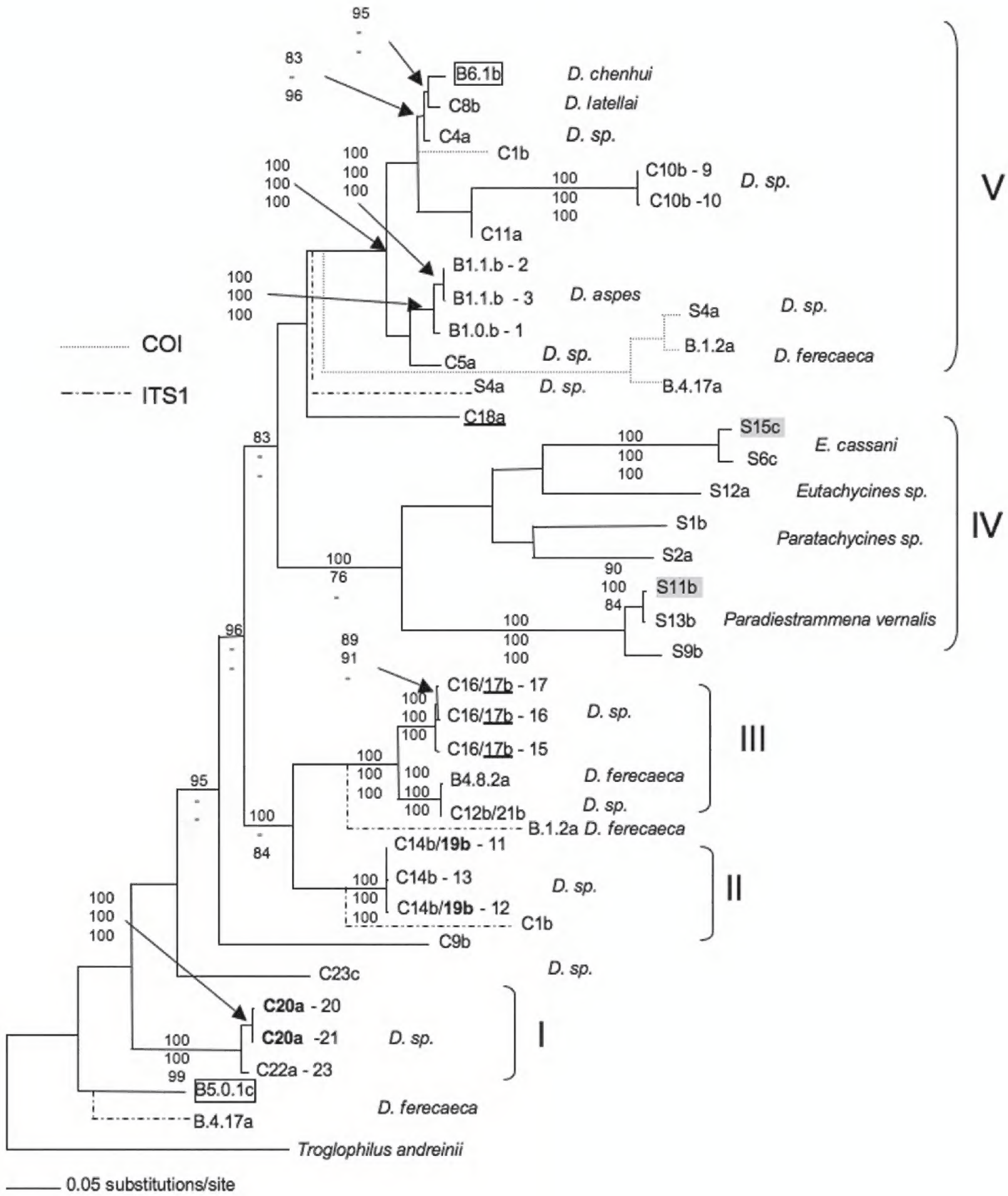

Fig. 4 - Evolutionary relationships among populations and species of Aemodogryllinae included in the study. The tree is a Bayesian phylogram; values at nodes are PP support for the Bayesian search (first value) and bootstrap supports for MP and NJ (second and third value; see text for details on the phylogenetic analyses). Only support values > 75% are shown. Underlined haplotypes were found in the Pi Pa Dong cave; bold haplotypes co-occur in the Xian Yen Dong cave; boxed haplotypes are from the An Ja Da Dong cave while haplotypes shaded in gray are from the Maria Cassan cave. Dotted (COI) and dashed (ITS1) branches represent cases of incongruence that could not be reconciled and are positioned in the tree according to the placements yielded by the phylogenetic searches conducted on the two genes separately. Five major clades are identified (see text for details).

that caused an increase in the amount of homoplasy. Finally, and perhaps more realistically, an incomplete taxon sampling should be considered as a likely reason. The cave cricket diversity in the area considered in this study is poorly known and new taxa are constantly being described (Rampini et al 2008; Zhang and Liu 2009). It is hence very likely that our taxon sampling is far from being complete and this could have affected the accuracy of our phylogenetic reconstruction.

Molecular systematics and phylogeography

Keeping in mind the limitations intrinsic to our study in terms of representativeness of taxa, a few considerations can nonetheless be done. The genus *Diestrammena*

(subgenus *Gymnaeta*) does not form a monophyletic cluster. The other genera included in the study (*Eutachycines, Paratachycines* and *Paradiestrammena*) are firmly clustered in clade IV, which is deeply embedded within the subgenus *Gymnaeta*. It shouldn't be overlooked that this is the first molecular study considering cave crickets from Laos and Thailand. COI-based K2P divergence within clade IV is 12.6% (SD = 4.1%) thus higher than the 7.64% K2P value below which Virgilio et al. (2010) could assign intraspecific comparisons in 95% of the ca. 16000 DNA barcodes they analyzed in insects. Within the subgenus *Gymnaeta* at least four clades (I-III and V in Fig. 4) could be identified molecularly; the only comparison close to the 7.64% K2P threshold is clade I vs. II (7.9% SD = 0.2%); the remaining comparisons always gave values higher than 8.3%. At the within clade level, clade V exceeds the 7.64% threshold (9.7% SD = 4.3%). Based on our data and on the above pieces of evidence, we conclude that the morphological characters identifying the subgenus *Gymnaeta* as such (Storozhenko 1991; Gorochov 1994) are questionable. More noticeably, the subgenus *Gymnaeta* as currently defined hosts a number of deeply divergent lineages whose systematics awaits proper consideration. At the species level, genetic data cluster all *D. (G.) aspes* haplotypes in a single clade; this is not the case with *D. (G.) ferecaeca*. This nominal taxon is distributed scatteringly in the tree of Fig. 4 and is one source of incongruence. According to Gorochov et al. (2006) *D. ferecaeca* exhibits a great deal of morphological variability such that different subspecies have been identified. We are inclined to suspect that such a morphological variability might conceal lineages more divergent than expected at the subspecific level. *D. (G.) chenhui* and *D. (G.) latellai* are genetically close; both species are from medium to large bodied and have well-developed eyes (Rampini et al 2008). In spite of these morphological similarities, it seems unlikely for the two species to be sister taxa. As a matter of fact *D. (G.) solida* and *D. (G.) zorzini* are morphologically and geographically closer respectively to *D. (G.) chenhui* and *D. (G.) latellai* than the latter are to one another. To fully resolve relationships among species we would have needed to screen molecularly *D. (G.) solida* and *D. (G.) zorzini* but we had no access to those species for our study.

The phylogenetic position of two of the four blind taxa (B5.0.1c and C23c), rather basal in the tree of Fig. 4, was not anticipated. Theoretically speaking, blind taxa are expected to end up in a given phylogenetic tree as terminal rather than basal branches with the latter being occupied by less cave-adapted lineages (Sbordoni et al 2000). This is because troglobionts are thought to derive from troglophile, trogloxene or even surface ancestors. An evolutionary scenario with a troglobitic taxon giving origin to a surface lineage is hard to envision, let alone this happening at least twice within the same group as suggested by the relationships in Fig. 4. We are more inclined to suspect that, once again, an incomplete taxon sampling is responsible for the observed pattern.

Fig. 5 is a schematic representation of genetic and geographic relationships among the five major clades of Fig. 4. It is evident that some clades identified phylogenetically are not restricted to a single area but they are spread across

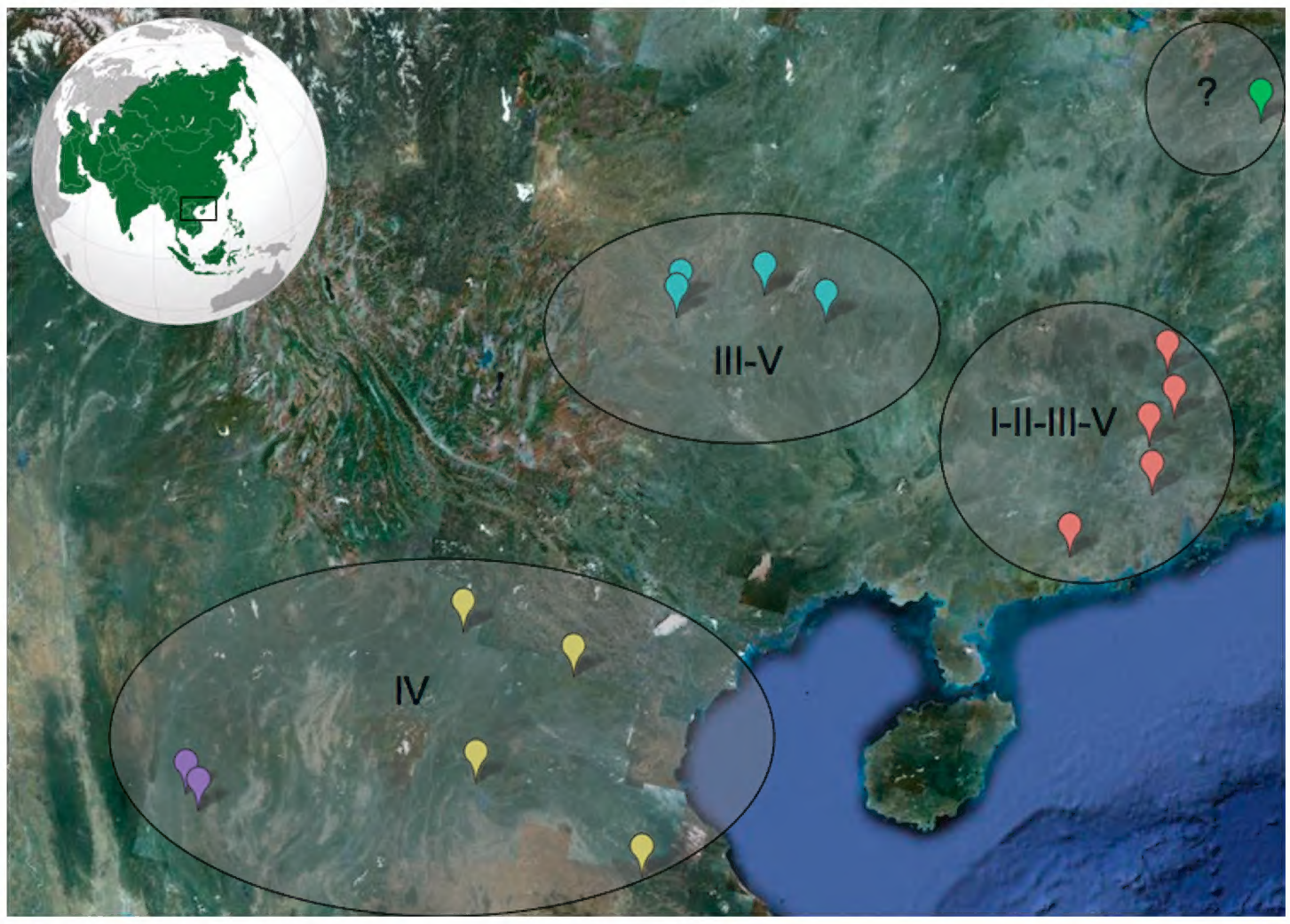

Fig. 5 - Geographic distribution of the main five clades identified phylogenetically in Aemodogryllinae representatives. Colors follow the same logic as in Fig.1.

two or more areas. Clades III and V co-occur in Guizhou (with a number of phylogenetically independent lineages) and Guangdong. This pattern is brought to an extreme in four caves (An Ja Da Dong, Pi Pa Dong, Maria Cassan and Xian Yen Dong), which host two divergent lineages each (see Tables 1 and 2 and Fig. 4 for details). It is noteworthy that lineages occurring within the same cave have different degrees of eye development (well developed vs. blind) suggesting a) a different timing in the colonization of the caves (older for the blind form, more recent for the form with developed eyes) and, b) divergence in the ecological niche, possibly related to the exploitation of different sections of the cave; this pattern has already been observed in *Diestrammena* (Di Russo and Rampini 2005).

All in all, the sympatric occurrence of genetically divergent lineages should be taken as an indication of secondary contact among lineages that have acquired their own characteristics allopatrically (Avise 2000). The alternating climatic conditions of the area during the Quaternary are likely to have strongly influenced the evolutionary history of these crickets. Although the area was never glaciated (Li et al 2007), throughout the Quaternary, phases of warm and humid climate alternated with colder and drier periods. Rhaphidophorids have quite strict requirements as far as environmental humidity is concerned because they cannot withstand dry conditions. It is thus reasonable to envision a scenario where the Quaternary dry and cold periods forced these organisms to seek refuge in the subterranean environment hindering surface gene flow. This led to local divergence and, if the time span was long enough, to the acquisition of troglobitic traits (Sbordoni et al 2000; Ketmaier et al 2002, 2010; Rampini et al 2008). The milder phases of the Quaternary (which partially still persists) allowed epigean dispersal and presumably brought into secondary contacts those lineages that had been evolving in allopatry. Such surface dispersal was limited to those groups that hadn't evolved a tight dependency on caves and would explain the occurrence within the same cave system of lineages divergent both genetically and in terms of the degree of troglomorphosis.

As a conclusive remark, we want to reiterate our awareness on the fact that this study has its own limitations geographically and in terms of coverage of taxa. Nonetheless, our genetic data have unveiled a complex and multifaceted evolutionary scenario that could be used as a backbone to thoroughly revise the taxonomic arrangements of the group in the area and as a starting point for more detailed research based on a larger number of populations and molecular markers.

ACKNOWLEDGMENTS

Many people involved in the biospeleological expeditions in Southern China (Alessandra Aspes, Daniele Avesani, Louis Deharveng, Leonardo Latella, Enrico Mezzanotte, Boris Sket, Helmut Steiner, Roberto Zorzin) have contributed material and are we are deeply thankful for having made this study possible. Financial support came from the MIUR (Italian Ministry of Instruction, University and Research) to MC and from the University of Potsdam (Germany) to VK.

REFERENCES

Alfaro, M. E., Zoller S., F. Lutzoni. 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Molecular Biology and Evolution 20: 255-266.

Allegrucci, G., Trewick, S. A., Fortunato, A., Carchini, G., V. Sbordoni. Cave Crickets and Cave Weta (Orthoptera, Rhaphidophoridae) from the Southern End of the World: A Molecular Phylogeny Test of Biogeographical Hypotheses. Journal of Orthoptera Research 19:121-130.

Allegrucci, G., Trucchi, E., V. Sbordoni. 2011. Tempo and mode of species diversification in *Dolichopoda* cave crickets (Orthoptera, Rhaphidophoridae). Molecular Phylogenetics and Evolution, 60: 108-121.

Avise, J. C. 2000. Phylogeography. The History and Formation of Species. Cambridge, London: Harvard University Press.

Clemet, M., Posada, D. K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9: 1657-1659.

Di Russo, C., M. Rampini. 2005. The cave crickets of Far East: a contribution to the study of East - Asian Rhaphidophoridae diversity (Orthoptera). Subterranean Biology 3: 37-42.

Farris, J. S. 1970. Methods for computing Wagner trees. Systematic Zoology 18: 374-385.

Farris, J. S., Kallersjo, M., Kluge, A.G., C. Bult. 1994. Testing significance of incongruence. Cladistics 10: 315-319.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. Journal of Molecular Evolution 17: 368-376.

Gorochov, A.V. 1994. New Orthopterans of the Infraorder Grasshoppers (Orthoptera, Tettigoniidae) from Vietnam and China. Proceedings Zoological Institute St. Petersburg 257: 18-50.

Gorochov, A.V. 2002. Material on the fauna and systematics of the Stenopelmatoidea (Orthoptera) of Indochina and some other territories. III. Entomologicheskoe Obozrenie 81: 318–337.

Gorochov, A.V., Rampini., M., C. Di Russo. 2006. New species of the genus *Diestrammena* (Orthoptera: Rhaphidophoridae: Aemodogryllinae) from caves of China. Russian Entomological Journal 15: 355-360.

Huelsenbeck, J.P. 2000. MrBayes: Bayesian inference of phylogeny. Distributed by the author. Department of Biology, Univ. of Rochester, Rochester, NY.

Ji, Y.-J., Zhang, D.-X. L.-J. He. 2003. Evolutionary conservation and versatility of a new set of primers for amplifying the ribosomal internal transcribed spacer regions in insects and other invertebrates. Molecular Ecology Notes 3: 581-585.

Ketmaier, V., De Matthaeis, E., M. Cobolli. 2002. Pattern of gene flow and genetic divergence in the three Italian species of the cave cricket genus *Troglophilus* (Orthoptera, Rhaphidophoridae): allozyme data. Subterranean Biology 2: 15-26.

Ketmaier, V., Di Russo, C., Rampini, M., M. Cobolli. 2010. Genetic divergence in the cave cricket *Troglophilus neglectus* (Orthoptera, Rhaphidophoridae): mitochondrial and nuclear data. Subterranean Biology 7: 25-33.

Ketmaier, V., Pirollo, D., De Matthaeis, E., Tiedemann, R., G. Mura. 2008. Large-scale mitochondrial phylogeography in the halophilic fairy shrimp *Phallocryptus spinosa* (Milne-Edwards, 1840) (Branchiopoda; Anostraca). Aquatic Sciences 60: 65-76.

Larget, B., D. Simon. 1999. Markov chain Monte Carlo algorithm for Bayesian analysis of phylogenetic trees. Molecular Biology and Evolution 16: 750-759.

Latella, L., R. Zorzin R. 2008. South China Research. Memorie del Museo Civico Di Storia Naturale di Verona -2. Serie. MonografieNaturalistiche 3.

Li, J., Ge, X. J., Cao, H. L., W. H. Ye. 2007. Chloroplast DNA diversity in *Castanopsis hystix* populations in south China. Forest Ecology and Management 243: 94-101.

Mau, B., M. Newton. 1997. Phylogenetic inference for binary data on dendrograms using Markov chain Monte Carlo. Journal of Computational and Graphical Statistics 6: 122-131.

Mau, B., Newton, M., B. Larget. 1999. Bayesian phylogenetic inference via Markov chain Monte Carlo methods. Biometrics 55: 1-12.

Posada, D., K. A. Crandall. 1998. Modeltest: testing the model of DNA substitutions. Bioinformatics 14: 817-818.

Rampini, M., Di Russo, C., M. Cobolli. 2008. The Aemodogryllinae cave crickets from Guizhou, Southern China (Orthoptera, Rhaphidophoridae). Pp. 129-142 in L. Latella, R. Zorzin, eds. South China Research, Memorie del Museo Civico Di Storia Naturale di Verona -2. Serie. Verona: Monografie Naturalistiche 3.

Rannala, B., H. Yang. 1996. Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. Journal of Molecular Evolution 43: 304-311.

Ronquist, F., J.P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.

Sbordoni, V., Allegrucci, G., D. Cesaroni. 2000. Population genetic structure, speciation and evolutionary rates in cave dwelling organisms. Pp. 453-477 in H. Wilkens, D. C. Culver, W. F. Humphreys, eds. Ecosystems of the World, vol 30. Subterranean ecosystems. Amsterdam: Elsevier.

Storozhenko, S.Y. 1991. Review of the subfamily Aemodogrylline (Orthoptera, Rhaphidophoridae). Entomologicheskoye Obozreniye 4: 810-813.

Swofford, D.L. 2002. PAUP* ß-version 4.10. Sinauer Associates Inc., Sunderland.

Thompson, J. D., Gibson,T. J., Plewniak, F., Jeanmougin, F., D. G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 24: 4876-4882.

Van der Niet, T., Lindner, H. P., Bytebier, B., D. U. Bellstedt. 2005. Molecular markers reject monophyly of the subgenera of *Satyrium* (Orchidaceae). Systematic Botany 30: 263-274.

Virgilio, M., Backeljau, T., Nevado, B., M. De Meyer. 2010. Comparative performances of DNA barcoding across insect orders. BMC Bioinformatics, 11: 206.

Zhang, F., X. W. Liu. 2009. A review of the subgenus *Diestrammena* (*Gymnaeta*) from China (Orthoptera: Rhaphidophoridae Aemodogryllinae). Zootaxa, 2272: 21-36.